



SNBU 2025

XXIII Seminário Nacional de
Bibliotecas Universitárias

17 A 20 DE NOVEMBRO
SÃO PAULO - SP

Eixo 5 – Acesso aberto

Representação da autoria indígena em repositórios institucionais: proposta de adaptação dos metadados na UFSCar

*Representation of Indigenous Authorship in Institutional Repositories: A Metadata
Adaptation Proposal at UFSCar*

Felipe Augusto Arakaki – Universidade de Brasília (UnB) – felipe.arakaki@unb.br

Caroline Periotto – Universidade Federal de São Carlos (UFSCAR) – carol@ufscar.br

Luciana de Souza Gracioso – Universidade Federal de São Carlos (UFSCAR) –
luciana@ufscar.br

Jair de Jesus Massa – Universidade Federal de São Carlos (UFSCAR) –
jair@estudante.ufscar.br

Luzia Sigoli Fernandes Costa – Universidade Federal de São Carlos (UFSCAR) –
luziasigoli@ufscar.br

Resumo: Apresenta uma proposta de adaptação dos metadados no Repositório Institucional da UFSCar com foco na inclusão de campos para autoidentificação étnico-racial e uso de línguas indígenas. Fundamentado nos princípios CARE e na curadoria de metadados, o estudo propõe novos campos e qualificadores no DSpace. Os testes demonstraram viabilidade técnica e compatibilidade com os fluxos existentes. A participação de bolsistas indígenas contribuiu para a legitimidade da proposta, que já se encontra em fase de implementação institucional. Os resultados apontam caminhos para práticas mais inclusivas e justas de representação informacional na ciência aberta.

Palavras-chave: Representação da informação. Metadados. Autoria indígena. Repositórios institucionais. Curadoria digital.

Abstract: This paper presents a metadata adaptation proposal for UFSCar's Institutional Repository, focusing on the inclusion of fields for ethnic self-identification and Indigenous languages. Based on the CARE principles and metadata curation, the study proposes new metadata fields and qualifiers in the DSpace platform. Tests





demonstrated technical feasibility and compatibility with existing workflows. The involvement of Indigenous fellows ensured the legitimacy of the proposal, which is already being implemented institutionally. The results suggest pathways for more inclusive and equitable information representation practices in the context of open science.

Keywords: Information representation. Metadata. Indigenous authorship. Institutional repositories. Digital curation.

1 INTRODUÇÃO

O avanço das políticas de ação afirmativa no ensino superior brasileiro, especialmente a partir da Lei nº 12.711/2012, promoveu o ingresso de estudantes indígenas em universidades públicas e estimulou o crescimento da produção acadêmica oriunda de epistemologias indígenas. Essa transformação quantitativa e qualitativa impõe novos desafios aos sistemas de informação, sobretudo no que tange à representação e visibilidade das identidades étnico-raciais e linguísticas nos repositórios institucionais. Apesar do reconhecimento constitucional das línguas e culturas indígenas (Brasil, 1988), a infraestrutura informacional das universidades, em grande parte, ainda opera com modelos descritivos universais que silenciam a diversidade dos modos de nomear, produzir e circular o conhecimento (Mignolo, 2003; Santos, 2010).

Neste cenário, o presente trabalho tem como objetivo apresentar uma proposta de adaptação dos esquemas de metadados do Repositório Institucional da UFSCar com vistas à inclusão de campos que permitam a autoidentificação étnica de autores indígenas, bem como a descrição de elementos textuais em línguas originárias. A iniciativa parte da premissa de que a produção acadêmica indígena demanda estruturas documentais sensíveis à sua especificidade, de modo a garantir o direito à representação autônoma e ao reconhecimento epistêmico nos sistemas de armazenamento e disseminação do conhecimento científico (Smith, 1999).

A justificativa do trabalho reside na necessidade de enfrentar as limitações dos atuais padrões de descrição, como o Dublin Core¹ e o MARC 21², que não contemplam

¹ O Dublin Core é um conjunto de quinze elementos genéricos e amplamente utilizados criado em 1995.

² MARC 21 (Machine-Readable Cataloging) é um formato de metadados para intercâmbio de dados bibliográficos e de autoridade que é legível por computador.



de forma explícita as categorias de etnia e língua indígena nos metadados bibliográficos. A proposta apresentada está fundamentada nos princípios CARE (*Collective Benefit* - Benefício Coletivo, *Authority to Control* - Autoridade para Controlar, *Responsibility* - Responsabilidade, *Ethics* - Ética)³, desenvolvidos pela *Global Indigenous Data Alliance* (2019), que orientam a governança ética e culturalmente adequada de dados indígenas.

Com base nesses fundamentos, este estudo visa contribuir para a construção de representações mais inclusivas e interoperáveis, capazes de acolher a diversidade epistêmica e linguística que compõem o cenário acadêmico brasileiro contemporâneo. Trata-se, assim, de uma iniciativa que integra inovação tecnológica, responsabilidade social e justiça cognitiva, por meio da disponibilização e curadoria de metadados e da participação ativa de sujeitos indígenas nos processos de representação informacional.

2 METODOLOGIA

A pesquisa adota uma abordagem aplicada e qualitativa, de natureza exploratória e descritiva, centrada na análise e modificação dos metadados utilizados no Repositório Institucional da UFSCar que opera com um *software* livre chamado *DSpace* versão 9.0. O estudo foi desenvolvido em três etapas articuladas: (1) levantamento e diagnóstico dos campos de metadados *Dublin Core* atualmente implementados no sistema; (2) revisão bibliográfica e normativa voltada à identificação de parâmetros técnicos e ético-políticos relacionados à governança de dados indígenas; e (3) proposição e teste de novos campos e qualificadores de metadados para representação da identidade étnico-racial e das línguas indígenas em produções acadêmicas.

A proposta metodológica se fundamenta na perspectiva da curadoria de metadados, conforme discutida por Pomerantz (2015) e Haynes (2018), e incorpora os princípios CARE (Global Indigenous Data Alliance, 2019) como orientação transversal. A partir dessa fundamentação, foi criado o campo específico `dc.contributor.authorethnicity`, para identificação da etnia ou povo indígena do(a) autor(a), e foram reconfigurados os campos `dc.title.alternative`, `dc.description.abstract`

³ Uma rede internacional que promove a soberania de dados indígenas, garantindo direitos, governança e participação dos povos indígenas para seu bem-estar e interesses coletivos.



e dc.subject, de modo a permitir a inserção de informações em línguas originárias, com codificação linguística baseada na norma ISO 639-2⁴.

A implementação foi realizada em ambiente de testes do DSpace, com validação técnica e documental. Paralelamente, a metodologia envolveu a participação de bolsistas indígenas da UFSCar, que contribuíram com sugestões terminológicas, avaliação da estrutura proposta e testes de submissão de documentos no repositório. Essa participação colaborativa foi fundamental para garantir a legitimidade da representação e o respeito à autodeterminação identitária dos sujeitos descritos.

Além da adequação técnica da estrutura de metadados, o estudo priorizou a escuta ativa e a construção coletiva, reconhecendo os limites dos padrões hegemônicos de descrição e buscando promover uma curadoria orientada por princípios de equidade, autonomia e reconhecimento. A metodologia adotada, portanto, conjuga análise técnico-informacional, revisão crítica de padrões e práticas de coautoria na representação da informação, reforçando o compromisso ético e epistêmico do projeto.

3 RESULTADOS E DISCUSSÕES

Os resultados alcançados pelo projeto expressam avanços na direção de uma infraestrutura de informação mais sensível à diversidade étnico-racial e linguística da produção científica contemporânea, especialmente no contexto dos repositórios institucionais mantidos por universidades públicas. A principal contribuição técnica consiste na extensão do esquema de metadados *Dublin Core* qualificado utilizado no *DSpace* para acolher atributos identitários e linguísticos de autores indígenas, historicamente invisibilizados nas estruturas convencionais de descrição bibliográfica.

Quadro 1 - Metadados para representação da língua e etnia indígena na produção científica da UFSCar.

⁴ A ISO 639-2 é um padrão internacional da ISO (Organização Internacional para Padronização) que define códigos de letras para representar nomes de idiomas e grupos de idiomas em todo o mundo.



Campo Dublin Core	Nome do Campo	Descrição do Campo
dc.contributor.authorethnicity	Etnia/Povo do autor (inclusão do campo)	Quando houver, informe a etnia/povo do autor (sem abreviações). Utilize letras maiúsculas apenas no início das palavras. Ex.: “Xukurú do Ororubá”. Você pode incluir mais de uma etnia/povo. Para isso, clique no botão “+ Adicionar mais”.
dc.contributor.author	Autor (adequação da descrição de preenchimento do campo)	Informe o nome completo do autor (sem abreviações). Utilize letras maiúsculas somente nas iniciais de nomes e sobrenomes. No caso de nome social ou nome indígena, utilize o nome como consta nos Sistemas UFSCar. Você pode incluir mais de um autor, para isso, clique no botão “+ Adicionar mais”.
dc.description.abstract	Resumo em outro idioma (adequação da descrição de preenchimento do campo)	Informe o resumo do trabalho acadêmico em outro idioma. Você pode incluir mais de um resumo em outro idioma. Se foi informado no campo anterior o resumo em português, informe neste campo o resumo em inglês, espanhol, língua indígena etc. Para isso, clique no botão “+ Adicionar mais”.
dc.subject	Palavras-chave (adequação da descrição de preenchimento do campo)	Informe as palavras-chave em português e em outros idiomas. Utilize letras maiúsculas apenas no início das palavras e em siglas. Para adicionar uma nova palavra-chave, clique no botão “+ Adicionar mais”.
dc.title.alternative	Título em outro idioma (adequação da descrição de preenchimento do campo)	Informe o título e o subtítulo em outro idioma. Se foi informado no campo anterior o título em português, informe neste campo o título em inglês, espanhol, língua indígena, etc. Você pode incluir mais de um título em outro idioma, para isso, clique no botão “+ Adicionar mais”.

Fonte: Dados da pesquisa (2025).



Descrição: Quadro com campos de metadados Dublin Core qualificado, adaptados para representar a etnia e a língua de autores indígenas na produção científica da UFSCar. Inclui campos para informar etnia, nome completo, resumo, palavras-chave e título em outros idiomas, com orientações de preenchimento que respeitam identidades culturais e linguísticas.

Foi criado e validado o campo `dc.contributor.authorethnicity`, permitindo o registro estruturado da etnia ou povo a que pertence o(a) autor(a) indígena, com base na autodeclaração. Este campo foi desenhado para manter coerência com os princípios da governança ética de dados (Global Indigenous Data Alliance, 2019) respeitando a soberania dos sujeitos sobre os modos de nomeação e descrição de si. A inclusão da etnia como elemento descritivo marca um deslocamento do paradigma universalista da representação, ancorado em neutralidade, para uma abordagem que reconhece o caráter situado, político e plural dos dados autorais. Esse metadado foi levado em consideração ainda, a partir de demandas de discentes indígenas da Universidade, com intuito de buscar e recuperar trabalhos de comunidades semelhantes e específicas desenvolvidas na própria Universidade.

Destaca-se ainda que antes de criar esse metadado, foi realizado um levantamento em bases de dados específicas de metadados como *Linked Open Vocabularies* (LOV) e o *MetadataRegistry*.

O *Linked Open Vocabularies* (LOV) - <https://lov.linkeddata.es/dataset/lov> - é uma iniciativa voltada à curadoria, catalogação e disseminação de vocabulários abertos utilizados na construção da Web Semântica. Trata-se de um repositório que reúne e descreve ontologias, esquemas de metadados e vocabulários controlados que seguem os princípios do Linked Data, permitindo sua reutilização e interligação em diferentes contextos informacionais. Desenvolvido originalmente pela comunidade *Open Knowledge Foundation Network* (OKFN) e atualmente mantido pela *Ontology Engineering Group* (OEG), o LOV fornece informações estruturadas sobre cada vocabulário, como prefixos, licenças, autores, alinhamentos com outros modelos e uso em datasets publicados. Além de promover a interoperabilidade semântica entre sistemas, o LOV funciona como um ponto de referência para desenvolvedores, pesquisadores e instituições que desejam adotar padrões abertos, promovendo boas



práticas na modelagem de dados conectados e contribuindo para a sustentabilidade de ecossistemas informacionais distribuídos (Vandenbussche et al., 2017).

Já o *Open Metadata Registry* (metadataregistry.org) configura-se como uma infraestrutura pública voltada à publicação, curadoria e compartilhamento de esquemas de metadados e vocabulários controlados, em conformidade com os princípios da Web Semântica⁵ e os padrões ISO/IEC 11179⁶. Sua função primordial é prover um ambiente confiável e persistente para o registro de elementos descritivos, conceitos codificados (value vocabularies) e ontologias, assegurando consistência terminológica e interoperabilidade entre sistemas heterogêneos. Cada vocabulário é descrito em termos de autoria, escopo, versão, URI persistente, alinhamentos semânticos e políticas de manutenção, o que possibilita sua reutilização em ambientes distribuídos, como repositórios digitais, plataformas de dados abertos e catálogos bibliográficos. O registro público dos esquemas favorece a transparência na construção dos metadados e estimula práticas colaborativas na governança da informação, sendo particularmente útil em contextos que exigem padronização, rastreabilidade e integração semântica de dados, como ocorre na construção de ecossistemas interoperáveis para a ciência aberta. (METADATA REGISTRY, 2010).

Adicionalmente, foram modificados os campos `dc.title.alternative`, `dc.description.abstract` e `dc.subject` para permitir o registro de elementos textuais em línguas indígenas, codificados segundo a norma ISO 639-2. Esta modificação torna possível que títulos, resumos e palavras-chave sejam inseridos em línguas como Nheengatu, Guarani, Tikuna, entre outras, assegurando visibilidade e recuperação informacional por meio de filtros linguísticos e abrindo espaço para a valorização da diversidade linguística no espaço acadêmico. Para tanto, foi realizado um levantamento

⁵ Web Semântica é uma extensão da Web tradicional que visa tornar o conteúdo das páginas compreensível não apenas por seres humanos, mas também por máquinas, por meio da padronização de dados e da atribuição de significados explícitos aos recursos da web. Ela se fundamenta no uso de tecnologias e linguagens como RDF (Resource Description Framework), OWL (Web Ontology Language) e SPARQL, que permitem a descrição formal de informações, suas relações e inferências lógicas. O objetivo central é facilitar a interoperabilidade, integração e reutilização de dados de maneira automatizada e inteligente, promovendo uma Web mais conectada, estruturada e orientada ao conhecimento.

⁶ ISO/IEC 11179 é uma norma internacional para a representação e registro de metadados (dados sobre dados). Ela define como descrever, registrar e gerenciar dados de forma padronizada, para que possam ser compreendidos, reutilizados e compartilhados entre diferentes sistemas, organizações e contextos.



das principais línguas indígenas identificadas no Brasil e que estão representadas na ISO 639-2. Dessa forma, é possível garantir a interoperabilidade entre outros sistemas de informação sem deixar de representar de forma adequada o resumo, assunto e títulos em línguas indígenas.

A viabilidade técnica foi confirmada por meio de um ambiente de testes configurado no DSpace 7.6.1, onde produções acadêmicas reais foram inseridas com os novos campos. A submissão de registros bilíngues (português–língua indígena) demonstrou estabilidade do sistema, compatibilidade com os fluxos de autoarquivamento existentes e possibilidade de retrocompatibilidade com os conjuntos de metadados anteriores.

Do ponto de vista metodológico e político, destaca-se a participação de bolsistas indígenas vinculados ao Programa de Educação Tutorial (PET) da UFSCar na concepção e validação da proposta. Esses estudantes contribuíram com sugestões terminológicas, discussão sobre autodeclaração, avaliação dos exemplos de preenchimento e identificação de barreiras culturais na usabilidade da plataforma. A experiência resultou em práticas de coautoria descritiva que tencionam os limites entre curadoria técnica e representação social, reforçando a ideia de que os sujeitos descritos devem participar das decisões sobre como são representados (Smith, 1999).

Trata-se da primeira iniciativa institucional no Brasil a operacionalizar a representação de autoria indígena por meio de campos de metadados específicos em um repositório acadêmico. Esse conjunto de resultados indica que a representação da autoria indígena em repositórios digitais não é apenas tecnicamente viável, mas desejável e urgente. Envolve o reconhecimento de saberes plurais, o respeito à diversidade cultural e o compromisso com uma ciência mais justa, aberta e inclusiva. A curadoria de metadados, nesse contexto, torna-se um instrumento de transformação social e de reparação histórica no campo da informação científica.

4 CONSIDERAÇÕES FINAIS

A experiência de ampliação dos metadados no repositório institucional da UFSCar para acolher elementos de auto-identificação étnico-racial e línguas indígenas demonstrou que os repositórios universitários possuem potencial para se tornarem



instrumentos de equidade epistêmica e reconhecimento da diversidade no ambiente acadêmico. Os resultados indicam que, com ajustes técnicos e com o envolvimento ativo dos sujeitos representados, é possível romper com estruturas descritivas homogêneas e promover práticas de representação mais inclusivas, culturalmente sensíveis e eticamente orientadas.

Entre os principais aprendizados, destaca-se a importância de incorporar a participação indígena de forma efetiva nas decisões relacionadas à representação da autoria. O diálogo entre equipes técnicas, profissionais da informação e estudantes indígenas foi essencial para garantir a pertinência dos termos adotados, a legitimidade da estrutura e a aceitabilidade do modelo proposto. Essa abordagem colaborativa se mostrou não apenas metodologicamente adequada, mas também politicamente necessária, reafirmando o princípio de que nenhuma representação pode ser neutra ou dissociada de seu contexto social e histórico.

No entanto, o projeto também enfrentou desafios. Entre as principais dificuldades, destacam-se: a limitação estrutural da plataforma DSpace em lidar com múltiplos idiomas em campos específicos sem customização; a ausência de vocabulários controlados amplamente reconhecidos para etnias indígenas; e a escassez de diretrizes nacionais voltadas à descrição de autoria indígena em sistemas informacionais. Esses aspectos revelam a urgência de um debate mais amplo sobre políticas de representação nos ambientes digitais da ciência aberta, especialmente aqueles voltados à produção de sujeitos historicamente marginalizados.

Como encaminhamentos futuros, propõe-se: (1) o desenvolvimento de vocabulários específicos e ontologias aplicadas à autoria indígena, em conformidade com os princípios FAIR⁷ e CARE; e (2) a criação de estratégias de formação continuada para profissionais da informação sobre representação sensível e antirracista nos acervos digitais.

⁷ Os princípios FAIR são um conjunto de diretrizes para a gestão e compartilhamento de dados científicos que tornam os dados Facilitadores de Busca (Findable), Acessíveis (Accessible), Interoperáveis (Interoperable) e Reutilizáveis (Reusable).



Por fim, ressalta-se que a curadoria de metadados constitui uma ação de mediação entre a técnica e o social, e pode ser mobilizada como ferramenta de reparação, justiça cognitiva e valorização de epistemologias plurais no espaço acadêmico. Ao incorporar campos que acolhem e valorizam as múltiplas formas de nomear, narrar e comunicar saberes indígenas, os repositórios deixam de ser apenas instrumentos de armazenamento e passam a constituir espaços de afirmação identitária, autonomia e transformação.

REFERÊNCIAS

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Senado Federal, 1988. Disponível em:

https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 30 jun. 2025.

BRASIL. **Lei nº 12.711, de 29 de agosto de 2012**. Dispõe sobre o ingresso nas universidades federais e nas instituições federais de ensino técnico de nível médio. *Diário Oficial da União*: seção 1, Brasília, DF, 30 ago. 2012.

GLOBAL INDIGENOUS DATA ALLIANCE (GIDA). **CARE Principles for Indigenous Data Governance**. 2019. Disponível em: <https://www.gida-global.org/care>. Acesso em: 30 jun. 2025.

HAYNES, David. **Metadata for information management and retrieval: understanding metadata and its use**. 2. ed. London: Facet Publishing, 2018.

METADATA REGISTRY. **About the Open Metadata Registry**. 2010. Disponível em: <http://metadataregistry.org/about.html>. Acesso em: 30 jun. 2025.

MIGNOLO, Walter D. **Histórias locais/projetos globais: colonialidade, saberes subalternos e pensamento liminar**. Tradução de Eliana Lourenço de Lima Reis; Wilma Freire da Silva. Belo Horizonte: UFMG, 2003. Disponível em: https://monoskop.org/images/7/7e/Mignolo_Walter_D_Historias_locais_projetos_globais_2003.pdf. Acesso em: 30 jun. 2025.

POMERANTZ, Jeffrey. **Metadata**. Cambridge, MA: MIT Press, 2015.

SANTOS, Boaventura de Sousa. Para além do pensamento abissal: das linhas globais a uma ecologia de saberes. In: SANTOS, Boaventura de Sousa; MENESES, Maria Paula (org.). **Epistemologias do Sul**. São Paulo: Cortez, 2010. p. 31–83.

SMITH, Linda Tuhiwai. **Decolonizing methodologies: research and indigenous peoples**. London: Zed Books, 1999. Disponível em: <https://nycstandswithstandingrock.wordpress.com/wp->



[content/uploads/2016/10/linda-tuhiwai-smith-decolonizing-methodologies-research-and-indigenous-peoples.pdf](#). Acesso em: 30 jun. 2025.

VANDENBUSSCHE, Pierre-Yves; ATEMEZING, Ghislain A.; POVEDA-VILLALÓN, María; VATANT, Bernard. Linked Open Vocabularies (LOV): a gateway to reusable semantic vocabularies on the Web. **Semantic Web**, v. 8, n. 3, p. 437–452, 2017. DOI: 10.3233/SW-160213. Disponível em: <https://www.semantic-web-journal.net/system/files/swj1178.pdf>. Acesso em: 30 jun. 2025.